ABSTRACT
The prediction of F'2 is an important aspect for vowel perception. Several prediction models have been proposed in the recent years. In these studies, relationships with the Center of Gravity (GC) in particular through the CRD on the vowel frequency were proposed in various models. Critical distance dynamic programming (CDP) and critical distance transformation (CDT) have been proposed as alternative methods. The evaluation results are carried out with two kinds of sound and give us a good spectral representation for a speech recognition system.

INTRODUCTION
Previous works have underlined two interesting phenomena: the center of gravity of spectral peaks and the F'2 of vowels /1-4/. Thus we can ask whether a second sound was added to the first sound. Can we predict F'2 by means of a similarity measure?

INTEGRATION
With these results, we can build an algorithm to find the best Fv frequency of a second sound and give us a good spectral representation for a speech recognition system.

EXPERIMENT
We found that it is hard to simulate both experiments with a machine "operator" instead of a listener. It is the aim of this paper to describe the machine operators we used.

CONCLUSION
The classic distance measures compare two spectra, component by component, at the frequency. If we draw a graph with Fv on the x-axis and F2 on the y-axis, we can find the path that is the closest to the diagonal. Some people proposed that any kind of path should be possible (Fv,F2). We used dynamic programming to get the best path. Each graph node (with coordinates x,y) had a weight which was computed by an elementary distance d(x,y).

ACKNOWLEDGMENTS
This research was supported by the French Ministry of Research and Technology, the French Ministry of National Education, and the European Community.
**Results with spectral peaks**

We use the results of a previous experiment which has been carried out essentially with nine swedish vowels (in Hz /1/):

\[
\begin{align*}
\text{Vowel} & \quad \text{Frequency} \\
\text{i} & \quad 310, 2350, 3300 \quad 730 \\
\text{e} & \quad 400, 710, 3650, 3150 \quad 720 \\
\text{o} & \quad 360, 1600, 2200, 3390 \quad 1720 \\
\text{a} & \quad 580, 940, 2440, 3990 \quad 960 \\
\text{u} & \quad 555, 1920, 2450, 3300 \quad 2050 \\
\text{U} & \quad 290, 1650, 2140, 3310 \quad 1720 \\
\text{e} & \quad 375, 2060, 2550, 3400 \quad 2350 \quad -0.05 \\
\text{a} & \quad 605, 1550, 2450, 3400 \quad 1860 \\
\text{u} & \quad 255, 2055, 2980, 3400 \quad 3210 \quad -0.01 \\
\end{align*}
\]

The estimated F2 by GDP and CDT (with peaks as inputs) are as following (in Hz):

\[
\begin{align*}
\text{F2} & \quad \text{GDP} & \quad \text{CDT} \\
\text{F2(CDP)} & \quad \text{F2(CDT)} & \quad \text{Eabs} \\
\text{u} & \quad 742 & \quad 740 & \quad 0.06 \\
\text{e} & \quad 725 & \quad 720 & \quad 0.01 \\
\text{o} & \quad 1830 & \quad 1888 & \quad 0.59 \\
\text{a} & \quad 949 & \quad 950 & \quad 0.09 \\
\text{u} & \quad 2084 & \quad 2200 & \quad 0.58 \\
\text{U} & \quad 1774 & \quad 1804 & \quad 0.01 \\
\text{e} & \quad 2216 & \quad 2340 & \quad -0.10 \\
\text{a} & \quad 1938 & \quad 1977 & \quad -0.01 \\
\text{u} & \quad 3097 & \quad 2980 & \quad -0.05 \\
\end{align*}
\]

Where E is the absolute error in Bark, Eabs is the total mean absolute error in Bark, \(E_{a} \) is minimal error.

In Fig.1 we can find a LPC spectrum and a CDT filtered spectrum of the same signal. This means that higher-spectral components are well integrated. With input FT spectra, the results are similar.

The correlation coefficient of CDT Euclidian distance with respect to human phonetic judgements are between 0.88(test X) and 0.85(test ABX) for the 11 french vowels. This means that CDT has retained a great deal of phonetic information. By comparison the Itakura distance obtained, with this method, the values 0.88(test X), 0.9(test ABX). As an example one can find on figure 3, the distance behaviour between two vowels.

\[ D(V_i, V') = \sum_{j=1}^{n} \frac{(F(j) - \mu_j)^2}{s_j} \]

\[ D(V_i, V') = \sum_{j=1}^{n} \frac{(F(j) - \mu_j)^2}{s_j} \]

CONCLUSIONS

We obtained also similar results with the center of gravity. But we found that our methods is possible to get good results for F2 or center of gravity. But it is very difficult to obtain good results for both of these experiments (F2 and CDT). Since some F2 values from previous experiments (Bladon & Carlson) seem incompatible and may be these values are also language depending. At last defined and are difficult to reproduce. For further work it is necessary to get more accurate values from human experiments.

**References**


**Fig.2a.** LPC spectrum of a /i/.

**Fig.2b.** CDT spectrum of the same /i/.

**Fig.3.** Two examples of distance behaviour (with a CDT preprocessing) between a vowel pairs (V and V') where X is a correlation coefficient between distances and perceptual data. Abx is an error number with respect to the perceptual boundary, t is an error number with respect to the perceptual boundary. r is a correlation coefficient between distances and perceptual data. Abx and Ax are two kinds of experiments. The zero crossing points are the discrimination points of the distance D. The arrows are human perception boundary.